

OPTIMIZACIÓN DE LAS SEÑALES ACÚSTICAS UTILIZANDO BEAMFORMING DE BANDA ANCHA PARA PERSONAS CON DISCAPACIDAD AUDITIVA

Paúl Bernal, Fabián Sáenz *, and Amelia Caisapanta

Departamento de Eléctrica y Electrónica DEEE,
Universidad de las Fuerzas Armadas - ESPE
Sangolquí Ecuador
{cpbernal, fgsaenzavcaisapanta}@espe.edu.ec
<http://www.espe.edu.ec>

Abstract. La voz es la principal forma de comunicación que posee el ser humano, el presente trabajo busca ayudar a las personas con problemas de audición, mediante el realzado de la voz. Para esto se utilizan las técnicas de *beamforming* y *Direction of Arrival* (DOA) en un arreglo de micrófonos semiesférico. Mediante el análisis de los algoritmos adaptativos *Linearly Constrained Minimum Variance* (LCMV) con sus tres familias: *Constrained*, *Generalized Sidelobe Canceler* (GSC) y *Householder*, utilizando los algoritmos *unconstrained: Least Mean Squares* (LMS), *Normalized Least Mean Squares* (NLMS), *Recursive Least-Squares* (RLS) y *Conjugate Gradient* (CG). Se pretende verificar las cualidades y desventajas de cada uno de los algoritmos tratando de optimizar sus desempeños variando sus respectivos parámetros de adaptación, a fin de obtener una rápida convergencia sin comprometer el *Mean Square Error* (MSE) ni aumentar el costo computacional. Los algoritmos adaptativos se comparan y en base a los resultados obtenidos se selecciona el algoritmo más idóneo, tomando en cuenta la velocidad de convergencia, costo computacional, todos los resultados son analizados a fin de obtener las conclusiones y recomendaciones.

Keywords: Mejoramiento de la voz, Beamforming, Direction of Arrival, Algoritmos Adaptativos, LCMV

1 Introducción

El procesamiento digital de señales de voz posee una amplia base teórica y experimental que se encuentra bien adaptada a las demandas computacionales del procesamiento digital de señales de voz [1]. El estudio de las señales de voz es un tema de importancia, dado que una gran parte de la sociedad presenta

* Estudiante del Doctorado en Ciencias Informáticas de la Universidad Nacional de La Plata, Argentina.

algún tipo de discapacidad auditiva, por tanto la presente trabajo pretende ser la base matemática para la implementación de soluciones que mejoren la calidad de vida de dichas personas.

Existen varias técnicas para el procesamiento de señales sonoras, dependiendo de la aplicación que se realice, en este caso se utiliza la metodología de *beam-forming* [2] y la estimación de la dirección de arribo de las señales (*Direction of Arrival*) DOA [3], que tradicionalmente eran estudiados en la teoría de antenas, substituyendo las ondas electromagnéticas por señales acústicas y el arreglo de antenas por un arreglo de micrófonos, presentando varias aplicaciones en diferentes campos. Los algoritmos son los encargados de realizar tales tareas, dentro de los cuales se puede nombrar: *Delay and Sum DAS*, Capon, *Multiple Signal Classification* MUSIC, Root-MUSIC, *Estimation of Signal Parameters via Rotational Invariance Technique* ESPRIT [4] y *Linearly Constrained Minimum Variance* LCMV [2].

De los algoritmos se pretende escoger uno que tenga buenos resultados y que su costo computacional sea razonable, para que después se implementen en un FPGA (*Field-Programmable Gate Array*). El presente artículo se encuentra compuesto de cinco secciones: En la Sección II se expone los conceptos básicos que se usan para la elaboración de este proyecto. En la Sección III, analiza el arreglo geométrico semiesférico. En la Sección IV se realiza un análisis de los resultados mas relevantes encontrados y en la Sección V se concluye el trabajo y se presentan trabajos futuros.

2 Conceptos Básicos

Para el desarrollo del presente artículo se inicio implementando un arreglo lineal ULA con un $M = 4$, esto pensado en futuras implementaciones con recursos de hardware y de economías limitadas. Como se indico en la Sección I, existe un sinnúmero de algoritmos que permiten solventar el problema como escoger el mas adecuada.

Para determinar el algoritmos a implementar se escogió como limitación el costo computacional, proyectado lo para una posterior implementación, utilizando el número de multiplicaciones, como medida, la cual se detalla en la Tabla 1 a continuación:

Table 1. Costo computacional de los algoritmos

Algoritmo	Número de Multiplicaciones
Delay and Sum	$\frac{180^\circ}{\Delta} (M^4 + 2M^3 + 4M^2 + 4M + 2)$
Capon	$\frac{180^\circ}{\Delta} (M^4 + 3M^3 + 4M^2 + 4M + 2)$
MUSIC	$\frac{180^\circ}{\Delta} (M^4 + 2M^3 + 7M^2 + 4M + 2)$

Donde M es el número de micrófonos y Δ° es la escala para el barrido del ángulo. Estos costos computacionales es un aproximado y está basado en cada uno de los códigos que se utilizan para estimar el DOA [4]. Como se observa el costo computacional es alto esto es debido a que existe inversión de matrices en los procesos. Los algoritmos LCMV no son presentados en la Tabla 1 puesto que tienen un gran cantidad de estructuras, que tienen costos computacionales variables; por lo cual son los más adecuados para su implementación, con resultados aceptables.

2.1 Algoritmos de Filtraje Adaptativo

Los algoritmos adaptativos intentan minimizar una función de costo $\xi[k]$ y que generalmente se relaciona con el error de salida [5]. En la Figura 1 se muestra la estructura de un *beamforming* adaptativo *broadband* con M sensores y N taps (número de coeficientes del filtro). La salida de la estructura de la Figura 1 se expresa como $y[k] = \mathbf{w}^H(k)\mathbf{x}(k)$:

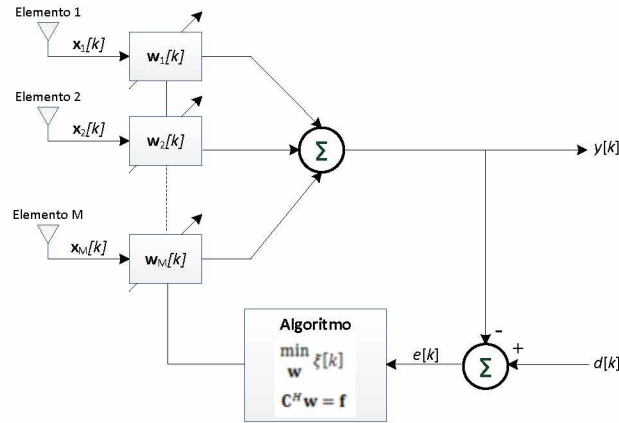


Fig. 1. Estructura de *Beamforming* Adaptativo *Broadband*.

$$\min_{\mathbf{w}} \xi(k) \quad \text{sujeto a} \quad \mathbf{C} = \mathbf{w}\mathbf{f} \quad (1)$$

donde $\xi[k] = E[|e[k]|^2]$. Con la matriz de restricción \mathbf{C} de dimensión $MN \times p$, siendo p el número de restricciones y \mathbf{f} el vector de ganancia de dimensión $p \times 1$ y corresponde a la ganancia unitaria en la fuente deseada y ganancia nula para las direcciones indeseadas, si fuera la aplicación a utilizar. Utilizando la función de costo y relacionando la señal de error $e[k] = d[k] - y[k]$, se obtiene la solución

de Wiener sujeto a $\mathbf{C}^H \mathbf{w} = \mathbf{f}$.

$$\mathbf{w}_{opt} = \mathbf{R}^{-1} \mathbf{p} + \mathbf{R}^{-1} \mathbf{C} (\mathbf{C}^H \mathbf{R}^{-1} \mathbf{C})^{-1} (\mathbf{f} - \mathbf{C}^H \mathbf{R}^{-1} \mathbf{p}) \quad (2)$$

Donde $\mathbf{R} = E[\mathbf{x}[k] \mathbf{x}^H[k]]$ y $\mathbf{p} = E[d^*[k] \mathbf{x}[k]]$ siendo \mathbf{R} es la matriz de autocorrelación de la señal de entrada y \mathbf{p} es el vector de correlación cruzada entre la señal deseada y la señal de entrada. Es fácil verificar que en ausencia de la señal deseada $d[k] = 0$ el vector $\mathbf{p} = \mathbf{0}$ y la solución óptima se reduce a la siguiente expresión:

$$\mathbf{w}_{opt} = \mathbf{R}^{-1} \mathbf{C} (\mathbf{C}^H \mathbf{R}^{-1} \mathbf{C})^{-1} \mathbf{f} \quad (3)$$

La solución de la Ecuación (3) es conocida como la ecuación óptima LCMV, donde la función de costo se denomina *Minimum Output Energy* (MOE). Los algoritmos adaptativos LCMV como ya se indicó requiere de un conjunto de restricciones que deben ser cumplidas en cada interacción. Mediante la imposición de restricciones lineales sobre el filtro adaptativo, la necesidad de una señal deseada a menudo puede ser descartada. Este tipo de algoritmos adaptativos se dividen en tres familias: con Restricciones Lineales (*Linearly Constrained*), de estructura GSC (*Generalized Sidelobe Canceler*) y los de estructura Householder [6].

2.2 Algoritmos Linearly Constrained

La actualización de los coeficientes en los algoritmos adaptativos *Constrained* se realiza en un subespacio que es ortogonal al subespacio generado por la matriz de restricción. La dirección de la actualización está dada por el vector de la señal de entrada premultiplicado por una matriz de proyección [7]. La matriz de restricción \mathbf{C} como se indicó es de dimensión $MN \times p$. Esta matriz tiene la forma siguiente:

$$\mathbf{C} = \begin{bmatrix} 1 & \dots & 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 1 & \dots & 1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 & 1 & \dots & 1 \end{bmatrix}^T \quad (4)$$

Al utilizar la matriz \mathbf{C} de la Ecuación (4) los algoritmos que se estudian en la presente investigación poseen un desempeño bajo, puesto que el valor del *Mean Square Error* MSE es alto, como se puede observar en la Figura 2, en la cual el valor promedio del MSE de cada uno de los algoritmos es de aproximadamente $-7dB$, por lo que se presume que al converger los algoritmos se encuentran en un mínimo local, puesto que a pesar de que se varía el tamaño del parámetro de adaptación de los algoritmos no se obtiene una disminución en el valor del MSE y un aumento exagerado en los valores de los parámetros de adaptación puede hacer que el algoritmo diverja. Para que los algoritmos mejoren su desempeño se preconditiona la estructura de la matriz \mathbf{C} , transformando esta matriz en una matriz toeplitz [8] y como se puede observar en la Figura 3 los algoritmos poseen un valor menor en el MSE en comparación a la Figura 2, con un valor promedio del MSE de $-28dB$ para CLMS, $-30dB$ para NLMS, $-33dB$ para CRLS y $-35dB$ para el algoritmo CCG.

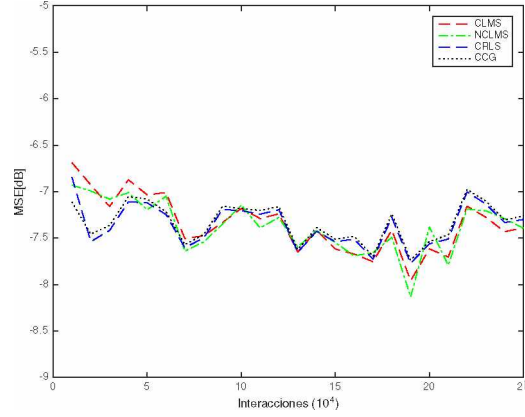


Fig. 2. Algoritmos Constrained para un arreglo ULA (\mathbf{C} original).

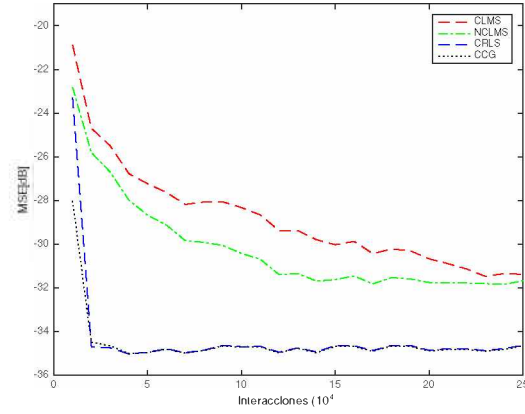


Fig. 3. Algoritmos Constrained para un arreglo ULA (\mathbf{C} modificada).

2.3 Estructura Generalized Sidelobe Canceler (GSC)

En la Figura 4 se presenta la estructura GSC. La estructura GSC es una alternativa diferente para implementar filtros adaptativos *linearly constrained*. En la estructura GSC, la dimensión del subespacio de adaptación se reduce adecuadamente a través de una matriz de *blocking* \mathbf{B} , permitiendo que cualquier algoritmo unconstrained pueda ser empleado como *constrained* [9].

Sea \mathbf{B} la matriz *blocking* de dimensión $MN \times (MN - p)$ que abarca el espacio nulo de la matriz de restricción \mathbf{C} . Existe una limitación para la matriz *blocking* y es que sus columnas deben ser ortogonales a la matriz de restricciones [6]. Una matriz de transformación \mathbf{T} de longitud $MN \times MN$, relaciona la estructura GSC con el filtro LCMV.

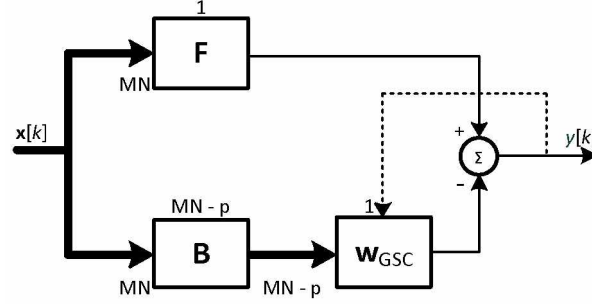


Fig. 4. Estructura Generalized Sidelobe Canceling (GSC).

$$\mathbf{T} = [\mathbf{C}^T \mathbf{B}] \quad (5)$$

La Figura 5 muestra el MSE de los algoritmo GSC, de los que se observa los algoritmos GSC-RLS y GSC-CG tienen el mejor desempeño

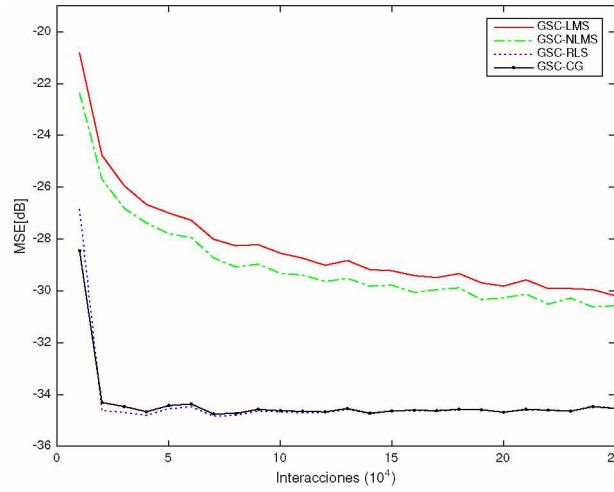


Fig. 5. MSE de los algoritmos GSC.

2.4 Transformada de Householder

La multiplicación de la matriz de *blocking* por el vector de señal de entrada en la estructura GSC puede ser computacionalmente alta y para muchas aplicaciones, no es práctico. La estructura Householder es una solución a este problema

mediante la matriz \mathbf{Q} y sus reflectores de Householder que cumple como la estructura GSC, pero computacionalmente son más eficientes [9].

La obtención del algoritmo HCLMS se inicia con la rotación en el vector $\mathbf{P}\mathbf{x}[k]$ con el fin de asegurar que el vector de coeficientes no se perturbe en otra dirección no excitado por $\mathbf{P}\mathbf{x}[k]$. Esto es posible si una matriz de rotación ortogonal \mathbf{Q} se utiliza como la transformación que va a generar un vector de coeficientes modificados $\hat{\mathbf{w}}[k]$ que se relaciona a $\mathbf{w}[k]$ [9].

$$\hat{\mathbf{w}}[k] = \mathbf{Q}\mathbf{w}[k] \quad (6)$$

La matriz \mathbf{Q} se debe escoger de tal manera que $\mathbf{Q}\mathbf{Q}^H = \mathbf{I}$. Los resultados de las simulaciones se observan en la Figura 6, destacando se nuevamente los algoritmos HCRLS y HCCG.

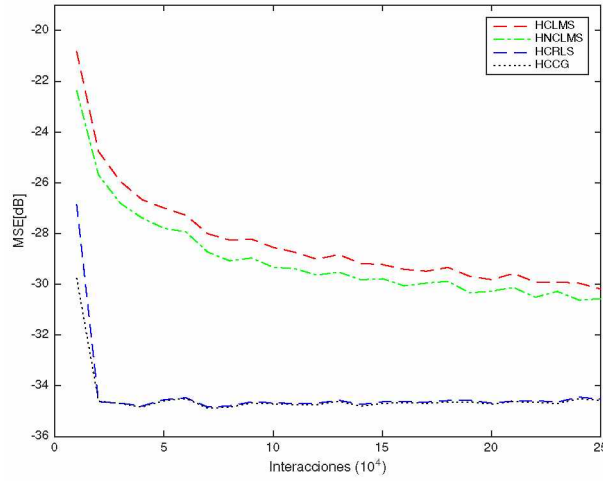


Fig. 6. MSE de los algoritmos Householder.

3 Análisis de los Algoritmos en la Geometría Semiesférica

En la literatura existen tres opciones para la configuración de un arreglo esférico [10]. La geometría que se utiliza en la presente investigación se basa en la configuración de L anillos con igual espaciamiento, donde los M_L micrófonos tienen una distancia de separación uniforme. El arreglo de micrófonos para la presente investigación consta de un semianillo $l = 1$ con 4 micrófonos espaciados uniformemente con una apertura entre cada micrófono adyacente de 60° , donde r es el radio con dimensión de 15 cm, posicionado en un plano (r_x, r_y) como se muestra en la Figura 7

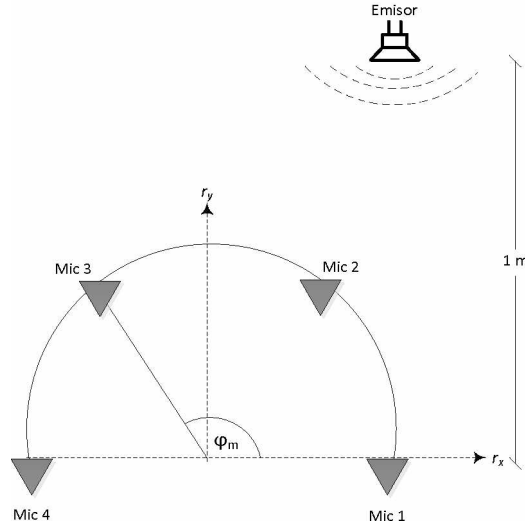


Fig. 7. Arreglo Semiesférico.

Las coordenadas esféricas de cada micrófono está dado por (r, θ, ϕ_m) . Donde $m = 0, \dots, 3$. Se utiliza cuatro micrófonos con $\theta = \pi/2$ y $\phi_m = m\pi/M$. Se supone un medio ideal, la velocidad de propagación de las ondas sonoras es constante y el tiempo de propagación depende solamente de la distancia recorrida. Para la presente investigación, el medio de interés es el aire cuyo valor es de $340m/s$ (no se considera las alteraciones debido a la altura en relación al nivel del mar). Según la Figura 7 la distancia d_m establece un camino directo entre la fuente emisora y el eje r_x del arreglo de los micrófonos. La distancia $d_m = 1m$.

$$t_m = \frac{d_m}{v} \quad (7)$$

El conocimiento del tiempo de propagación del emisor a los micrófonos constituye la base para el cálculo de los diferentes tiempos de llegada ya que la señal llega a los micrófonos en distintos tiempos, por tanto el tiempo de propagación desde la fuente emisora a los micrófonos Mic 1 y Mic 4 será el siguiente: $t_1 = t_4 = 1m/(340m/s) = 2,94ms$ El tiempo de propagación desde la fuente emisora a los micrófonos Mic 2 y Mic 3 es: $t_2 = t_3 = 0,87m/(340m/s) = 2,56ms$. Entonces, la señal llega al micrófono Mic 1 con un retardo de tiempo de 0,38 ms con relación al micrófono Mic 2 y de manera similar sucede con la señal que arriba al micrófono Mic 4 que lo hace con un retardo de 0,38 ms con relación al micrófono Mic 3. Las señales recibidas por los micrófonos son muestreadas a una tasa $f = 22050Hz$.

4 Análisis de los Resultados

Para realizar el análisis de los resultados partimos del costo computacional, en relación al número de multiplicaciones como podemos observar Tabla 2, la estructura que menos multiplicaciones emplea es la estructura de Householder, como ya se mencionó los reflectores householder mejoran el costo computacional.

Table 2. Costo computacional de los algoritmos LCMV

Alg.	Constrained	GSC	Householder
LMS	$N^2 + (2p + 2)N + 1$	$N^2 + (3 - p)N - (2p - 1)$	$(2p + 2)N - (p^2 - 1)$
NLMS	$N^2 + (3p + 3)N + 1$	$N^2 + (4 - p)N - (3p - 1)$	$(2p + 3)N - (p^2 + p - 1)$
RLS	$6N^2 + (19p + 4)N + p$	$5N^2 + (6 - 9p)N + 4p^2 - 5p$	$4N^2 + (5 - 6p)N + 2p^2 - 3p + 2$
CG	$4N^2 + (4p + 10)N + 1$	$4N^2 + (11 - 7p)N + 3p^2 - 10p + 1$	$3N^2 + (10 - 4p)N + p^2 - 8p + 2$

Otra manera de verificar los resultados es mediante la gráfica *beam pattern* (patrón de directividad) de la Figura 8, de cada uno de los algoritmos Householder, según dicha figura los algoritmos presentan una buena resolución en sus *beam patterns* y el ángulo obtenido coincide con el de la señal de entrada de 90° .

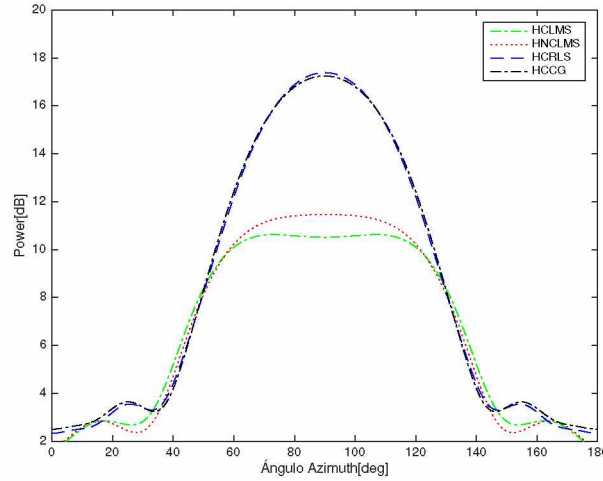


Fig. 8. Algoritmos Householder respecto al *beam pattern*

Los algoritmos HCRLS y HCCG poseen un similar comportamiento, por tanto una mayor ganancia, de aproximadamente 6 *dB* más que los algoritmos HCLMS y HNCLMS.

5 Conclusiones

En el presente trabajo corrobora, que las estructuras LCMV, son las de menor costo computacional sobre las técnicas tradicionales de *beamforming* y *Direction of Arrival* (DOA) empleadas; a cambio provee una solución razonable con un costo computacional bajo y un MSE pequeño. En cuanto a la familia *constrained*, se necesita condicionar mejor la matriz **C**, lo que conlleva un aumento en el costo computacional, por lo cual la estructura idónea mas es la estructura Householder, ya que sus reflectores son optimizados versus la estructura GSC y su costo computacional es reducido como se indica en Tabla 2. De esta estructura el algoritmo HCCG, es el que presenta mayores ventajas respecto a los otros, como se comprobo en el análisis de los resultados.

En trabajos a futuro, se pretende encontrar la forma de condicionar la matriz **C**, de modo a obtener mejores resultados, con un MSE menor, sin comprometer el costo computacional, y con la idea de una implementación en un FPGA.

References

1. Rabiner L., & R. W., S. Theory and Application of Digital Speech Processing, 2009.
2. Van Veen, B., & Buckley, K. Beamforming Techniques for Spatial Filtering. CRC-netBASE, 2000 .
3. Monzingo R., Haupt R., & Miller, T. Introduction to Adaptive Arrays (Second ed.). Published by SciTech Publishing, Inc. ,2011.
4. W. Y., Cho, Y. S., & Choo, C. Y. (2012). MATLAB /Simulink for Digital Signal Processing. Seoul, Korea: Hongrunc Publishing Company.
5. Paulo S. R. Diniz, Adaptative Filtering Algorithms and Practical Implementation, Third Edition, Springer, 2008
6. S. Chandran, Adaptive Antenna Arrays Trends and Applications. Springer-Verlag Berlin Heidelberg, 2004.
7. M. de Campos, S. Werner, & J. Apolinário. Constrained Adaptation Algorithms Employing Householder Transformation. IEEE Transactions on Signal Processing Vol. 50, No. 9, 9. September 2002
8. K. Gundersen, & J. Hakon Husoy, . Preconditioner structures for the CLMS adaptive filtering algorithm. IEEE. 2006
9. S. Werner, J. Apolinário, & Marcello L. R. , d. (s.d.). On the Equivalence of the Constrained RLS and the GSC-RLS Beamformers. Helsinki University of Technology, Instituto Militar de Engenharia, and Universidade Federal do Rio de Janeiro.
10. Z. Huang & C. Balanis, Adaptive Beamforming Using Spherical Array. IEEE, 2005.